

# Wireless Sensor Network Simulation for Fault Detection in Industrial Processes

Rui Pinto<sup>1</sup>, Rosaldo J. F. Rossetti<sup>1,2</sup> and Gil Gonçalves<sup>1</sup>

<sup>1</sup>*Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal*

<sup>2</sup>*Artificial Intelligence and Computer Science Lab, University of Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal*  
{rpinto, rossetti, gil}@fe.up.pt

**Keywords:** Castalia, Sensor Diagnosis and Validation, Industrial Wireless Sensor Networks.

**Abstract:** Sensor data is extremely important to monitor machines at the shop-floor level and its environmental surrounding conditions for condition-based monitoring, machine diagnosis and process adaptation to new requirements. Based on the described scope, self-diagnostics and self-organizing capabilities are core functionalities of any Industrial Wireless Sensor Network (IWSN). In the present work, a simulated case study was developed with the main intent of validating techniques implemented for sensor data diagnosis of error detection and equipment failure. The scenarios explored try to mimic some common situations of a manufacturing environment when dealing with WSNs, where a piece of sensor equipment suddenly stops working or an unpredictable change in the environment leads to faulty data readings. This paper introduces Castalia and describes how it was used to simulate a direct application of an Optical Metrology System on an industrial Resistance Spot Welding process, which is composed of a camera and several luminosity sensors. More specifically, a sensor data validation module was proposed, implemented and used to extend Castalia functionalities.

## 1 INTRODUCTION

Nowadays, several European initiatives towards smart manufacturing systems and Industry 4.0 are a current subject of research, aiming at the improvement of the European Industry competitiveness regarding other leading markets. This can be achieved by turning shop-floor production methods more flexible and efficient facing the constant changes in production, which is driven by consumer demands. Since consumers' preferences are becoming more customized and competitiveness between companies is extremely high, manufacturers must quickly adapt their product to new trends. Usually, the company pioneer on having a new product available on the market have the upper hand over the competition. When a mass production model is used, maintaining a large envelope of products or enlarging it with more product variants is extremely difficult and entails a great effort and cost due to the inflexibility of mass production manufacturing systems. To support the mass customization paradigm at the shop-floor level, the production systems should be capable of rapidly reconfiguring as product demand varies, as well as does machine parameter calibration during all production stages. The

time spent to calibrate and to adjust the machine process parameters when a new product variation should be performed, or after a maintenance phase, can be reduced if the machines' parameter adjustment is done automatically instead of the usual trial-and-error approach, by performing destructive tests to evaluate the process quality. Monitoring machine's execution is extremely important and, to do so, sensors are deployed on the shop-floor to infer surrounding environmental conditions and adapt the machine execution accordingly.

An Optical Metrology System (OMS) is a system where optical measurements are used for several industrial applications, e.g. for robot guidance when optical metrology sensors are mounted on a robot arm. In this case, a OMS is used on a Resistance Spot Welding (RSW) process, to infer if the metal parts to be weld are well positioned. In order to insure the correct setting up and easy ramping up of the system, in site factory sensor positioning adjustment and image setup is a significant procedure, which involves configuration of camera parameters, such as exposure and gain. With help of luminosity sensors by measuring the luminosity conditions in site, the camera parameters can be adjusted on the fly, according to

the context. This work focus on the simulation of an OMS applied to a RSW process using Castalia. The goal was to understand the behaviour of the luminosity sensors, specially when they are malfunctioning. A sensor data detection module was implemented in Castalia and used on a simulated version of the described environment.

The paper is organized in four more different sections. Section 2 details the approaches regarding IWSN, the main methods for sensor data validation and categorize WSN simulators, specially Castalia. Section 3 identifies the case study used to develop the simulation scenario. Section 4 depicts all the results obtained, whereas in Section 5 final remarks about the overall work exposed in the paper and further steps to be followed are presented.

## 2 INDUSTRIAL WIRELESS SENSOR NETWORKS

In the past few years WSNs have become more interesting to be explored in and applied to several domains. Their use was motivated mainly due to latest advances in wireless communications as well as in more reliable, robust and long-lasting hardware. These factors have a great impact on the feasibility of installation, when it is difficult to use wired solutions, either by harsh location or high number of sensor nodes used and also because of the easy maintenance and cabling reduced costs. As main advantages, Chen et al. (Chen et al., 2015) pinpoints the large coverage area, ubiquitous information, fast communication via RF and self-organisation throughout the direct communication between entities. A Smart Sensor Platform (Ramamurthy et al., 2007) was developed, which applies the plug and play concept by means of hardware interface, payload, communication between sensors and actuators, and ultimately allows for software update using 'over-the-air' programming. Cao et al. (Cao et al., 2008) developed a distributed approach to put closer sensors and actuators in a collaborative environment using WSNs and, Chen et al. (Chen et al., 2015), push this approach forward considering the same methodology, but taking into account all the industrial domain restrictions.

Despite the huge potential of WSNs, these solutions still have several issues and future research challenges. Data collected from WSNs is prone to be faulty due to internal and external influences, such as environmental effects, hardware malfunctions, software problems, energy constraints, network issues, security threats, among others, as shown elsewhere (Tolle et al., 2005), (Barrenetxea et al.,

2008), (Ramanathan et al., 2006) and (Szewczyk et al., 2004). IWSNs are used to monitor real-time production equipment in the factory and the surrounding environmental conditions, acting accordingly when changes are observed. According to Neumann (Neumann, 2007), there are important restrictions in industrial applications to be met, such as real-timeliness, functional safety, security, energy efficiency, and Quality of Service (QoS). So, IWSNs face some challenges, such as safety-critical functions, security and privacy of collected data, availability to avoid complete production stop when failures occur, latency/retransmission of messages, support for actuators (by using the same sensor controller), efficient integration with existing automation infrastructures, scalability, coexistence and wireless communication interference avoidance and energy harvesting (Åkerberg et al., 2011).

### 2.1 Sensor Data Validation

To maintain QoS, one should be able to detect and deal with compromised sensor nodes during operation. To determine whether a sensor node is malfunctioning, validation methods are applied to sensor data, aiming to find devious sensor readings from normal ones. There is no such an ideal validation method, because they are very dependent on sensor measurement conditions and the overall environment context (Freitas et al., 2010), (Branisavljević et al., 2011), (Bertrand-Krajewski et al., 2003), (Vasconcelos et al., 2012), and (Liu et al., 2013). Usually, several methods are applied successively to the sensor data, because each method is suitable to detect particular types of faulty data. Sensor data validation methods are typically divided into online and offline methods, where online methods are efficient, simpler to implement and demand less data processing, and offline techniques, such as Bayesian Networks, Artificial Neural Networks and Regression Techniques, are more robust and complex techniques, which are suitable to process data not in real-time.

The most common WSNs' data fault types are: Out-of-range faults - values out of predefined minimum and maximum boundaries; Struck-at-fault sequence of values that have little or no variation for a large period of time; Outliers misplaced data samples that represent a sensor misbehaviour, which can be caused by sensor malfunctioning or changes in surrounding environment; Stop Communication gaps in the dataset caused by an absence of sensor information on the gateway, due to power outage or communication problems; Devious Data when the behaviour of a certain sensor is devious from the majority, and it

persists for a large period of time. A graphical example of these faults' behaviour (on a temperature sensor) is presented in Figure 1.

According to Sharma et al. (Sharma et al., 2010), Ravichandran and Arulappan (Ravichandran and Arulappan, 2013), the most common techniques used to perform online detection of the described sensor data faults are Min/Max Detection, Flat Line Detection, Modified Z-Score, No value Detection and Spatial Correlation. The Min/Max Detection finds out-of-range faults using upper and lower limits that are defined by process limitations. On the other hand, the Modified Z-Score finds outliers faults which, despite being inside the process limitations, are values that are unusual when compared to the sensor data history. The Flat Line Detection finds struck-at-faults, which are characterized by little or no variation in data and the No value Detection detects situations when the data stops being communicated. Finally, the Spatial Correlation is the technique used to compare datasets of several sensors nodes physically located near to one another, and if one of them presents a devious dataset in comparison to its neighbours, then the corresponding sensor node is probably faulty.

## 2.2 WSN Simulation

A WSN simulation consists in using simulators specifically designed to imitate WSNs behaviour. Traditionally, the three main methodologies for analysing the performance of WSNs are physical measurements, which consists in setting up all the physical nodes and collecting sensor data, analytical methods, and computer simulation. Since deploying a WSNs requires a huge effort, due to constraints regarding the cost associated with hardware and the fact that analytical models are not effective, because of the complexity of the models for WSNs regarding energy limitation, decentralized collaboration and fault tolerance, simulation is the only feasible approach to the quantitative analysis of sensor networks, getting access to fine grained results easier than with real world experiments (Dwivedi et al., 2010). Because there are so many simulators with many different characteristics, Eriksson (Eriksson, 2009) proposes to categorize WSNs as Generic Network Simulators - focus more on network aspects, such as radio protocols, network stacks and channel distortions; Code Level Simulators - focus more on the simulation of the sensor nodes, such as deployable code and functionality logic; Firmware Level Simulators - focus more on emulating the sensor node, such as microprocessor, radio chip and other peripherals.

Song (Song et al., 2011) implemented a fault de-

tection mechanism to analyse the stability and reliability of data transmission in a ZigBee network, which was simulated on the NS-2 simulator. Results focus on the data loss rate in the communication between sensor nodes. Dai (Dai et al., 2011) also proposed a fault detection mechanism, but focused more on the delays introduced by the wireless communication networks, where the control system was simulated in MATLAB and the network-induced delays were simulated in OMNeT++. Both Song and Dai focused more on faults with root cause on network aspects. On the other hand, Zhang (Zhang et al., 2009) and Szczodrak (Szczodrak et al., 2008) used Castalia to implement a fault detection algorithm based on temporal and spatial correlations of sensor data from neighbour sensor nodes, classifying each node as good or faulty.

### 2.2.1 Castalia

Castalia (Boulis, 2007) was the simulator chosen for this work. Figure 2 presents Castalia's architecture, where each module accepts messages from other modules or itself and, according to the message, it executes a given code. The nodes do not connect to each other directly, but through the Wireless Channel module, which estimate the average path loss between two nodes using the Lognormal Shadowing. The nodes are also linked through Physical Process modules that they monitor, which "feed" the sensors with data. For every physical process there is one module which holds the "truth" on the quantity of the physical process that is representing. The nodes sample the physical process in space and time to get their sensor readings.

The Physical Process modules are based on an arbitrary number of point sources whose 'influence' is diffused over space and they can change their position and their value over time. The effect of multiple sources in a certain point is additive. Calculating the value of the physical process at a certain location and at a certain time is represented in Equation (1), where  $V(p, t)$  denotes the value of the physical process at point  $p$  and at time  $t$ ,  $V_i(t)$  denotes the value of the  $i^{th}$  source at time  $t$ ,  $d_i(p, t)$  denotes the distance of point  $p$  from the  $i^{th}$  source at time  $t$ ,  $K$  and  $a$  are parameters that determine how the value from a source is diffused.  $N(0, \theta)$  is a zero-mean Gaussian random variable with standard deviation  $\theta$ . The "ground truth" offered by the physical process is distorted by the inaccuracies of the sensing devices, implemented by the Sensor Manager.

$$V(p, t) = \sum_i \frac{V_i(t)}{(Kd_i(p, t) + 1)^a} + N(0, \theta) \quad (1)$$

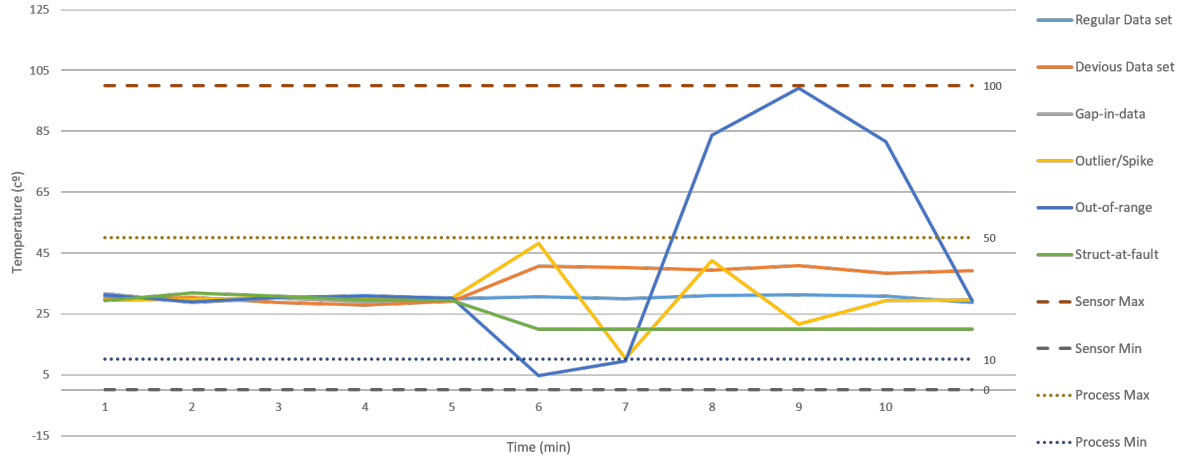


Figure 1: Sensor Data Faults Characterization.

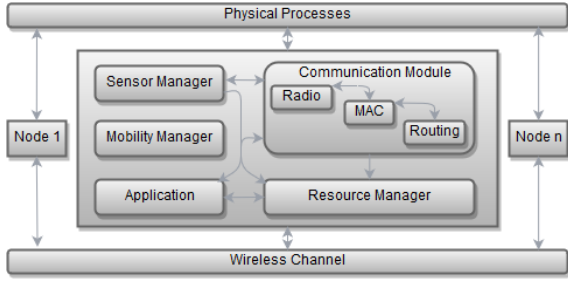


Figure 2: Model Structure of Castalia.

Regarding the internal structure of a node, the Application module is the one that specifies the algorithm that receives sensor data and acts accordingly. The MAC and Routing modules define several communication protocols, which can be used by the nodes. The Mobility module defines the nodes position and mobility pattern. The Sensor Manager defines the sensing devices present in each node, by introducing a distortion of the ground truth offered by the physical process. The Resource Manager defines the energy consumed by the different components of the node.

### 3 METHODOLOGY

The simulation scenario tried to mimic a OMS, which is composed by an IWSN, where all nodes are equipped with luminosity sensors and a central node is equipped with a camera. This OMS is used to analyse the quality of a RSW process, namely if both parts to be weld are correctly placed and, in the end of the welding cycle, evaluates the quality of the weld. This analysis consists in capturing and processing a camera image captured, which must have a minimum quality,

otherwise, image analysis will be impossible. In order to avoid excessive darkness/clarity in the image, the exposure index of the camera must be calibrated according to the luminosity conditions in the surrounding environment. This process consists in controlling the lightness and darkness of the image, which can be done by the camera light meter. Since the camera light meter may have a lower accuracy and, it measures only one single point of luminosity per sample, several luminosity sensors are used to infer the luminosity conditions around the welding area and improve the control of the camera exposure index.

In this case, a camera is placed right on top of the two electrodes that will weld the parts together. In order to evaluate the luminosity conditions, nine luminosity sensors were placed around the electrodes. All the sensor data is processed on a gateway node, resulting on an average of the nine luminosity points. The simulation of this scenario will allow to implement and test a sensor fault detection module before being deployed in a real environment. Figure 3 represents a real application of this scenario and the schema of the virtual world to be simulated. There are eight sensor nodes (represented by the green circles), each one equipped with a luminosity sensor that outputs values between 0 and 100%. At the center of the world are placed the luminosity source and sensor node 0. Even numbered sensor nodes are placed at 20 cm distance around the center, so their output should be similar. The same is true for odd numbered sensor nodes, which are placed at 50 cm distance and, in consequence, their output should be smaller than the even numbered sensor nodes.

Regarding the simulation scenario, only one luminosity source was considered (placed at the center of the world), which had default parametrization values,

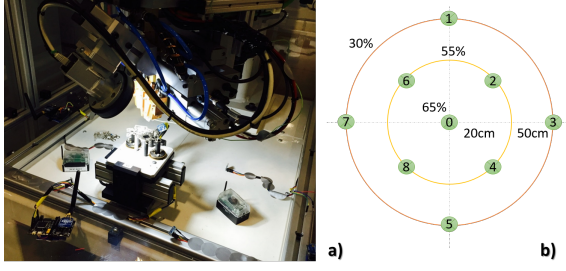


Figure 3: OMS used on a RSW process: a) real prototype; b) simulation.

namely  $K$  equals 0.25,  $a$  equals 1.0 and  $\theta$  equals 0.2. The output of the source was defined at a theoretical value of 65%. According to the simulated world implemented, sensor node 0 outputs higher luminosity values (around 65%), because it is the one closest to the luminosity source. Then, even numbered sensor nodes output values around 55% and odd numbered sensor nodes output values around 30%. For the communication module, the radio's parameters used are the ones proposed in the IEEE 802.15 Task Group 6 documents, the routing protocol is the Multipath Ring Routing and the MAC protocol is the IEEE 802.15.4.

## 4 TESTS AND PRELIMINARY RESULTS

Some tests were performed in order to simulate deviations in sensor readings. The sensor manager's device noise property was changed accordingly, which is by default as low as possible (but larger than 0). When this value increases, the difference of sensor output values compared to normal outputs also increases. Bigger deviations resulted on out-of-range and outlier faults. On the other hand, a device noise equals to 0 corresponds to a null variation, resulting on struck-at-fault samples. Figure 4 plots the percentage of simulated errors in the dataset of sensor node 0 and the efficiency of the techniques for error detection, calculated by determining the false positives and false negatives.

The considered dataset had 18% of out-of-range faults, 36% of outliers and spikes and 14% of struck-at-faults. The remaining 32% of the dataset is normal data. The Min/Max Detection pinpointed all out-of-range faults, presenting zero false positive and negative. The Flat Line Detection technique missed one out of 14 struck-at-faults present in the data set, presenting 1 false positive and zero false negatives. The Modified Z-Score detected only 12 out of 18 outliers present in the sensor data set, having 6 false positives and zero false negatives. The Modified Z-Score tech-

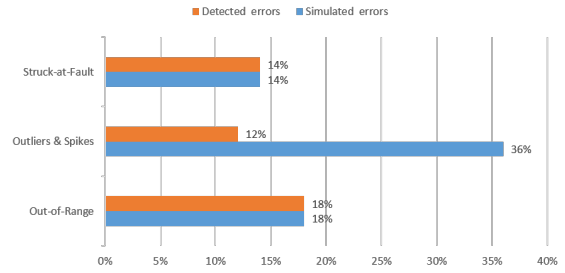


Figure 4: Efficiency of techniques for sensor data error detection.

niques efficiency depends greatly on the data set history, namely the rate between the number of normal samples and the number of outliers, and the difference between the normal and outlier sample values. In order to improve the efficiency of this method, the number of normal samples must be much larger than the number of outliers.

## 5 CONCLUSIONS AND FUTURE WORK

The present work simulates an OMS applied to an industrial RSW process, which analyses the position of materials to be welded by means of a camera and several luminosity sensors. The simulation implements an IWSN whose nodes have a luminosity sensor. In Castalia, sensors are "fed" by a luminosity physical process and run a sensor data validation module to detect failing sensors. Such a sensor data validation module implements Min/Max Detection, Flat Line Detection and Modified Z-Score, which are used to detect out-of-range faults, struck-at-faults, and outliers and spikes, respectively. All techniques performed as expected, even though the Modified Z-Score accuracy demonstrated to depend greatly on the history of dataset samples, as well as on the mean difference between normal and outlier samples as discussed in the result analysis.

Regarding future steps, the sensor data validation module can be improved by implementing the Spatial Correlation method, considering complex scenarios with different location luminosity sources and detecting the presence and location of a sensor fault. This complex feature implies the use of online sensor fusion techniques of several heterogeneous sensor data or offline complex methods, such as machine learning techniques. Such an approach to using multiple data types and sources is the essence of a multi-agent methodology (Passos et al., 2011) and architecture (Rossetti et al., 2007) currently under development,

also with applications to the industrial domain (Braga et al., 2008).

## REFERENCES

- Åkerberg, J., Gidlund, M., and Björkman, M. (2011). Future research challenges in wireless sensor and actuator networks targeting industrial automation. In *Industrial Informatics (INDIN)*, pages 410–415. IEEE.
- Barrenetxea, G., Ingelrest, F., Schaefer, G., Vetterli, M., Couach, O., and Parlange, M. (2008). Sensorscope: Out-of-the-box environmental monitoring. In *Information Processing in Sensor Networks (IPSN'08)*, pages 332–343. IEEE.
- Bertrand-Krajewski, J.-L., Bardin, J.-P., Mourad, M., and Beranger, Y. (2003). Accounting for sensor calibration, concentration heterogeneity, measurement and sampling uncertainties in monitoring urban drainage systems. *Water Science & Technology*, 47:95–102.
- Boulis, A. (2007). Castalia: revealing pitfalls in designing distributed algorithms in wsn. In *5th International Conference on Embedded networked sensor systems*, pages 407–408. ACM.
- Braga, R., Rossetti, R., Reis, L. P., and Oliveira, E. C. (2008). Applying multi-agent systems to simulate dynamic control in flexible manufacturing scenarios. In *Cybernetics and Systems 2008*, pages 488–493, Vienna. OSGK.
- Branišavljević, N., Kapelan, Z., and Prodanović, D. (2011). Improved real-time data anomaly detection using context classification. *Journal of Hydroinformatics*, 13(3):307–323.
- Cao, X., Chen, J., Xiao, Y., and Sun, Y. (2008). Distributed collaborative control using wireless sensor and actuator networks. In *Future Generation Communication and Networking (FGCN'08)*, pages 3–6. IEEE.
- Chen, J., Li, H., Sheng, D., and Li, W. (2015). A hybrid data-driven modeling method on sensor condition monitoring and fault diagnosis for power plants. *International Journal of Electrical Power & Energy Systems*, 71:274–284.
- Dai, X., Long, Z., Yang, B., and Tan, X. (2011). Robust fault detection observer in wireless networked control systems. In *Chinese Control and Decision Conference (CCDC)*, pages 2147–2152. IEEE.
- Dwivedi, A., Patle, V., and Vyas, O. (2010). Investigation on effectiveness of simulation results for wireless sensor networks. *Information Processing and Management*, pages 202–208.
- Eriksson, J. (2009). Detailed simulation of heterogeneous wireless sensor networks. Licentiate thesis, Uppsala University, Sweden.
- Freitas, T. R., Coelho, A., and Rossetti, R. (2010). Correcting routing information through GPS data processing. In *Intelligent Transportation Systems (ITSC)*, pages 706–711. IEEE.
- Liu, S., Araujo, M., Brunskill, E., Rossetti, R., Barros, J., and Krishnan, R. (2013). Understanding sequential decisions via inverse reinforcement learning. In *Mobile Data Management (MDM)*, pages 177–186. IEEE.
- Neumann, P. (2007). Communication in industrial automation what is going on? *Control Engineering Practice*, 15(11):1332–1347.
- Passos, L. S., Rossetti, R., and Gabriel, J. (2011). An agent methodology for processes, the environment, and services. In *Intelligent Transportation Systems (ITSC)*, pages 2124–2129. IEEE.
- Ramamurthy, H., Prabhu, B., Gadh, R., and Madni, A. M. (2007). Wireless industrial monitoring and control using a smart sensor platform. *Sensors Journal, IEEE*, 7(5):611–618.
- Ramanathan, N., Balzano, L., Burt, M., Estrin, D., Harmon, T., Harvey, C., Jay, J., Kohler, E., Rothenberg, S., and Srivastava, M. (2006). Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks. *Center for Embedded Network Sensing*.
- Ravichandran, J. and Arulappan, A. I. (2013). Data validation algorithm for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 2013.
- Rossetti, R., Oliveira, E. C., and Bazzan, A. (2007). Towards a specification of a framework for sustainable transportation analysis. In *13th Portuguese Conference on Artificial Intelligence, EPIA, Guimarães, Portugal*, pages 179–190. APPIA.
- Sharma, A. B., Golubchik, L., and Govindan, R. (2010). Sensor faults: Detection methods and prevalence in real-world datasets. *ACM Transactions on Sensor Networks (TOSN)*, 6(3):23.
- Song, C.-C., Feng, C.-F., Wang, C.-H., and Liaw, D.-C. (2011). Simulation and experimental analysis of a zigbee sensor network with fault detection and reconfiguration mechanism. In *Asian Control Conference (ASCC)*, pages 659–664. IEEE.
- Szczodrak, M., Zahedi, S., Ji, P., Mylaraswamy, D., Srivastava, M., and Young, R. (2008). Simulation framework for qoi characterization of sensor networks in the presence of faults.
- Szewczyk, R., Mainwaring, A., Polastre, J., Anderson, J., and Culler, D. (2004). An analysis of a large scale habitat monitoring application. In *2nd International Conference on Embedded networked sensor systems*, pages 214–226. ACM.
- Tolle, G., Polastre, J., Szewczyk, R., Culler, D., Turner, N., Tu, K., Burgess, S., Dawson, T., Buonadonna, P., Gay, D., et al. (2005). A macroscope in the redwoods. In *3rd International Conference on Embedded networked sensor systems*, pages 51–63. ACM.
- Vasconcelos, G., Petry, M., Almeida, J. E., Rossetti, R., and Coelho, A. L. (2012). Using UWB for human trajectory extraction. In *24th European Modeling and Simulation Symposium, EMSS*, pages 428–433.
- Zhang, C., Ren, J., Gao, C., Yan, Z., and Li, L. (2009). Sensor fault detection in wireless sensor networks. In *IET International Communication Conference on Wireless Mobile and Computing*, pages 66–69.